Estatística Inferencial

Prof. Wagner Hugo Bonat

Departamento de Estatística
Universidade Federal do Paraná

Teoria assintótica do EMV

Normalidade assintótica da função escore

- Principal resultado

$$\mathrm{U}(\theta|\boldsymbol{Y}) \overset{a}{\sim} \mathrm{N}(0, \mathrm{I}_{\mathrm{E}}(\theta)),$$

  onde

$$\mathrm{U}(\theta|\boldsymbol{Y}) = \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f(\theta, Y_i).$$

- A função escore é a soma de v.a iid para um dado θ. Pelas igualdades de Bartlet, temos

$$\mathrm{E}(\mathrm{U}(\theta|\boldsymbol{Y})) = 0 \quad \mathrm{e} \quad \mathrm{Var}(\mathrm{U}(\theta|\boldsymbol{Y})).$$

- Pelo Teorema Central do Limite, temos

$$\frac{\mathrm{U}(\theta|\boldsymbol{Y}) - \mathrm{E}(\mathrm{U}(\theta|\boldsymbol{Y}))}{\sqrt{\mathrm{Var}(\mathrm{U}(\theta|\boldsymbol{Y}))}} \overset{a}{\sim} N(0, 1).$$

Estimador de máxima verossimilhança (EMV)

- Estimativa de máxima verossimilhança: Seja $L(\theta, \mathbf{y})$ a função de verossimilhança. O valor $\hat{\theta} = \hat{\theta}(\mathbf{y})$ é a estimativa de máxima verossimilhança para θ se $L(\hat{\theta}) \geq L(\theta), \forall \theta$.

- Estimador de máxima verossimilhança: Se $\hat{\theta}(\mathbf{y})$ é a estimativa de máxima verossimilhança, então $\hat{\theta}(\mathbf{Y})$ é o estimador de máxima verossimilhança (EMV).

- Em muitos casos $\hat{\theta}$ é um máximo local no interior de Θ e satisfaz

$$U(\hat{\theta}|\mathbf{Y}) = 0.$$

Propriedades do EMV

- $\hat{\theta}$ é consistente, ou seja,
$$\hat{\theta} \xrightarrow{P} \theta, \quad \text{quando} \quad n \rightarrow \infty.$$
- Isso significa que
$$P(|\hat{\theta} - \theta| \geq \epsilon) \rightarrow 0 \quad \text{quando} \quad n \rightarrow \infty.$$
- $\hat{\theta}$ é assintóticamente normal e eficiente, ou seja,
$$\hat{\theta} \overset{a}{\sim} N(\theta, I_E(\theta)^{-1}).$$
- $\hat{\theta}$ é o melhor estimador disponível para θ quando o tamanho da amostra é grande.
- Assintóticamente não-viciado e eficiente.
- Exemplos com dados simulados (`ScriptPoisson.R`).

Método delta

- Se uma sequência Y_n satisfaz

$$\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2) \quad \text{quando} \quad n \to \infty,$$

e se $g : \Re \to \Re$ diferenciável em θ e $g'(\theta) \neq 0$, então

$$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2 g'(\theta)^2) \quad \text{quando} \quad n \to \infty.$$

- Demonstração.

Consistência: Caso geral

- Lembre-se do Teorema que deu origen a ideia do EMV.

$$P(l(\theta_0) > l(\theta)) \to 1 \quad \text{quando} \quad n \to \infty.$$

  para $\theta \neq \theta_0$ fixado.

- Teorema (Consistência): Com probabilidade tendendo a 1 quando $n \to \infty$, a verossimilhança tem uma solução $\hat{\theta}$ que é consistente.

- Demonstração (opcional ver Notas adicionais).

Eficiência e normalidade assintótica

- Assumimos que existe uma função $M(y)$ tal que

$$\left|\frac{\partial^3}{\partial\theta^3}\log f(y;\theta)\right| < M(y),$$

e $\mathrm{E}(\mathrm{M}(y)) < \infty$. Então

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} \mathrm{N}(0, \mathrm{i}_E(\theta)^{-1}) \quad \text{quando} \quad n \to \infty.$$

onde $i_E(\theta)$ é a informação de Fisher para uma observação.

- Demonstração.

Algumas implicações

- Qualquer termo assintóticamente equivalente a $\mathrm{I_E}(\theta)$ pode ser usado. Assim,

$$\hat{\theta} \sim NM_p(\theta, \mathrm{I_E}^{-1}(\hat{\theta}))$$

$$\hat{\theta} \sim NM_p(\theta, \mathrm{I_O}^{-1}(\theta))$$

$$\hat{\theta} \sim NM_p(\theta, \mathrm{I_O}^{-1}(\hat{\theta})).$$

- **Distribuição assintótica da deviance** - Para um problema de estimação regular, no limite com $n \to \infty$, se θ é o verdadeiro valor do parâmetro, então

$$\mathrm{D}(\theta) = -2[\mathrm{l}(\theta) - \mathrm{l}(\hat{\theta})] \sim \chi^2_d$$

ou seja, a função deviance segue uma distribuição qui-Quadrado com p graus de liberdade, onde p é a dimensão do vetor θ.

Resumo dos resultados

- O EMV $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ é assintóticamente não-viciado, isto é, $\mathrm{E}(\hat{\boldsymbol{\theta}}) \to \boldsymbol{\theta}$.
- Assintóticamente $\mathrm{V}(\hat{\boldsymbol{\theta}}) \to \mathrm{I_E}^{-1}(\boldsymbol{\theta})$, mostrando que o EMV é eficiente para o vetor $\boldsymbol{\theta}$, ao menos para grandes amostras.
- Denote $\mathrm{J} = \mathrm{I_E}^{-1}(\boldsymbol{\theta})$, então $\mathrm{V}(\hat{\boldsymbol{\theta}}) = \mathrm{J}$, sendo que, J é uma matriz simétrica e definida positiva, com elementos $\mathrm{J_{ij}} = \mathrm{Cov}(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_j)$ então $\mathrm{J_{ii}}$ é a variância de $\hat{\boldsymbol{\theta}}_i$.
- Podemos construir intervalos de $100(1-\alpha)\%$ de confiança para θ_i na forma $\hat{\theta}_i \pm z_{\frac{\alpha}{2}} \mathrm{J_{ii}}^{\frac{1}{2}}$.

Família exponencial de distribuições (Tópico adicional)

- Uma v.a Y é dita ser da familia exponencial se sua fp ou fdp tem a seguinte forma

$$f(y;\theta) = a(y)\exp^{\theta y - \kappa(\theta)},$$

  onde θ é chamado parâmetro canônico e $\kappa(\cdot)$ é uma função conhecida.
- Teorema: $\mathrm{E}(Y) = \kappa'(\theta)$ e $\mathrm{Var}(Y) = \kappa''(\theta)$.
- Demonstração.

Propriedades do EMV na familia exponencial

- Se Y_i são v.a iid com distribuição de probabilidade pertencente a familia exponencial temos:
  - O EMV $\hat{\theta}$ é consistente para θ.
  - O EMV é assintóticamente eficiente e normalmente distribuído.
- Demonstração (ver Notas adicionais).

Exemplos

- Sejam $Y_i \sim G(\theta)$ iid para $i = 1, \ldots, n$.
  - Encontro o EMV.
  - Qual a distribuição assintótica do EMV neste caso? Especifique sua esperança e variância.
- Sejam $Y_i \sim B(p)$ iid para $i = 1, \ldots, n$.
  - Encontro o EMV.
  - Qual a distribuição assintótica do EMV neste caso? Especifique sua esperança e variância.
- Sejam $Y_i \sim P(\lambda)$ iid para $i = 1, \ldots, n$.
  - Encontro o EMV.
  - Qual a distribuição assintótica do EMV neste caso? Especifique sua esperança e variância.